The Relationship Between Body Length and Scale Length in Five Year-Classes of the Pacific Pilchard or Sardine, *Sardinops caerulea* (Girard, 1854).

ANTONIO LANDA[1]

INTRODUCTION

BY CORRELATING the marginal growth of scales (the amount of growth between the last ring or annulus and the margin of the scale) with the season of the year in which the scales were collected, Walford and Mosher (1943*a*: 9 and 1943*b*: 12) have shown for the Pacific pilchard or sardine, *Sardinops caerulea* (Girard) 1854, that these rings are formed annually and consequently can be used for age determination as well as for back-calculation of the length of the fish at a given earlier age. The application of the scale rings to the problem of age determination in the pilchard has been thoroughly discussed in the papers cited above. The application of the same phenomenon to back-calculation of lengths, which lengths are used in several different types of problems, requires further consideration. The literature on this subject is voluminous and of considerable interest. Reference should be made to the works of Blackburn (1949), Buchanan-Wollaston (1934), Lea (1910, 1913), Lee (1912), Ottestad (1938), and Schuck (1949) for further information.

The estimation of body length at a previous age, using length of the scale from its center to the ring corresponding to that age, involves the regression of body length on scale length. The regression of scale length on body length, sometimes used, will give the mean scale length corresponding to a given body length, information which is not pertinent to the problem. It should be noted at once that different regression lines could be obtained from a given set of material depending on the manner of sampling and that this manner, in turn, depends on the use for which the regression values are intended. Thus, if the goal is to know at which body length the scales start growing, it should be emphasized that the sampling be comprehensive of the very small body lengths. Moreover, such sampling should be equally representative of all the body lengths if a further goal is desired, i.e., to know whether or not the regression of body length on scale length is the same throughout all different body lengths.

In the particular case of the Pacific pilchard investigations, and, presumably, in others of similar commercial fisheries, the goal is to know the body lengths that the fish caught commercially had in previous years. It follows that the sampling should, before anything else, be representative of the commercial catch even if that requirement makes the data unsuitable for the attainment of goals of the type indicated above. Once the scope of the problem is thus limited, the following are some of the questions that arise in the back-calculation and use of estimated lenghts:

(1) Are the regression constants (regression coefficient and body intercept) of body length on scale length the same for all year-classes regardless of region of capture?

(2) Are the regression constants for a given year-class the same regardless of region of capture?

[1] Biologist, Cia. Admora del Guano, Lima, Peru.
Manuscript received December 4, 1951.

(3) Can the regression of body length on scale length be expressed by a straight line when fish of successive ages are considered?

(4) If the regression is linear, does it have its origin at 0.0? (If so, the actual process of calculating lengths at previous ages may be greatly simplified.)

Answers to these four questions are necessary in order to decide if, for back-calculation of lengths, a given body of data should be treated as a unit or grouped by year-classes, by regions in which the fish were captured, by age groups, or by groups combining two or three of these categories.

Acknowledgments

I wish to thank Mr. John C. Marr, Dr. Frances E. Felin, and Mr. Theodore M. Widrig of the U. S. Fish and Wildlife Service for making the material used in this study available to me and for other assistance.

SOURCE OF DATA

Material with which to investigate these questions consists of records of scale readings for the Pacific sardine on file at the laboratory of the South Pacific Fishery Investigations, U. S. Fish and Wildlife Service, at Stanford University. The material was obtained as a result of a co-operative program between the U. S. Fish and Wildlife Service, the California Department of Fish and Game, the Fisheries Research Board of Canada, the Washington State Department of Fisheries, and the Fish Commission of Oregon. The methods of collecting samples, examining scales, and recording data are explained by Felin and Phillips (1948). In brief, each fish is represented by a card which contains information on fish length, sex, date and locality of capture, and, on a printed scale, the relative position of the focus, annuli, and margin of the fish scale. The data are, in general, representative of the populations commercially fished along the Pacific Coast of the United States and Canada (with the exception of bait fisheries in southern California).

From the material available (about 60,000 cards), the data for the 1938 through 1942 classes for fish caught off Canada, Washington and Oregon, San Francisco, Monterey, and San Pedro during the fishing seasons of 1939–40 through 1947–48 were selected in order to have (1) a series of year-classes represented, (2) all the commercial ports represented (3) for each year-class, fish with one to five annuli represented.

METHODS

To reduce this material (about 40,000 cards) to workable proportions and still have a representation of the fish taken by the commercial fishery, a sample was taken in the following manner:

(1) All the cards belonging to the material selected were considered as a single "population" without regard to year-class, port of landing, age, or season of the year in which the fish were caught.

(2) One thousand one hundred and ninety-nine bibliographic cards (about 3 per cent of the above population) were marked with numbers taken from a random numbers table (Table 1 in Dixon and Massey, 1949) using five-digit numbers. The cards were arranged in ascending order. Three percent was chosen because, after some trials, it was seen to be the minimum number of cards likely to include every category in the sample.

(3) The scale reading cards were counted, and, when the count coincided with the number on the first bibliographic card, the latter was removed, and the corresponding scale card was taken as part of the sample. This process was repeated, without interrupting the original count, until all the bibliographic cards were removed.

(4) The sample so obtained was then subdivided into all possible year-class and port combinations. By chance, the

TABLE 1

CODED VALUES TAKEN FROM SCATTER DIAGRAMS
(Y = Body Length; X = Scale Length)

YEAR-CLASS AND PORT COMBINATION	nj*	Y′†	Y′²	X′‡	X′²	X′Y′
1938 in:						
Pacific Northwest...	10	−29	1015	14	88	91
San Francisco..	91	−16	4514	79	843	1138
Monterey......	76	269	4371	26	640	938
San Pedro.....	60	−26	1692	−32	374	549
1939 in:						
Pacific Northwest...	43	138	5472	−33	349	807
San Francisco..	150	1079	27721	16	2302	5388
Monterey......	166	640	15716	−63	1541	2620
San Pedro.....	124	−128	4926	54	716	1065
1940 in:						
Pacific Northwest...	14	25	1681	22	45	96
San Francisco..	50	−12	2120	−4	368	430
Monterey......	75	487	12303	93	1111	2639
San Pedro......	70	−109	3823	−23	535	1091
1941 in:						
San Francisco..	5	−4	102	3	37	15
Monterey......	26	125	3171	42	274	726
San Pedro.....	90	153	2923	90	484	809
1942 in:						
Pacific Northwest...	7	61	1415	6	84	270
San Francisco..	23	16	1531	6	192	204
Monterey......	68	491	12327	159	1283	3464
San Pedro.....	51	−6	4486	10	502	1308

*nj = number of items in jth group.
†Y′ = sum of coded body lengths.
‡ X′= sum of coded scale lengths.

1941 year-class in the Pacific Northwest did not appear in the sample.

For each year-class and port combination a two-variable frequency table was made, plotting standard body lengths against scale lengths. Body lengths were grouped by 2-millimeter intervals, scale lengths by 6-millimeter intervals. In order to code, arbitrary means were independently chosen for each of the variables in every frequency table. Owing to the magnification used in reading the scales, the scale-length values are 30 times larger than the actual scale lengths (Felin and Phillips, 1948).

The values obtained from these frequency tables are given in Table 1. It is possible to pool these values in such a way as to obtain combinations other than port and year-class combinations. For example, to obtain information about the 1938 class as a unit, the sums of X′ for the 1938 class from all ports are pooled in a single sum, the sums of Y′ are similarly pooled, etc.

TABLE 2

VALUES USED IN THE COVARIANCE TESTS

TEST	S_1	S_2	S_3	S_4	N*	p†
1	53941	1046	646	125	1199	5
2	6631	26	882	1162	237	4
3	21828	292	1522	18	483	4
4	9154	392	421	2351	209	4
5	2899	175	30	405	121	3
6	5542	401	279	322	149	4

* N = total number of observations.
† p = number of groups.

Pooled data obtained in this manner were used to get the S_1, S_2, S_3, and S_4 values (see Table 2) as defined in Kendall (1946: 238 ff.) to make the tests indicated in Table 3. Tests were made after Kendall (*op. cit.*), at the .05 level of significance and each consisted of testing the hypotheses (1) that the regression coefficients of the subclasses considered could have been drawn from the same populations and (2) that the regression of body length on scale length was a straight line (see footnote to Table 3). They were performed with coded data; later the data were uncoded to calculate the means, deviations, regression coefficients of body length on scale length, and body-length intercepts for the groups that were shown, by the tests, to be significantly different. Test 1 indicated that the hypothesis (1) above should be rejected, and therefore it was necessary to consider each year-class separately.

RESULTS

From the results of Tests 1 to 6 (Table 3) and from the values found (Table 4), the questions posed can be answered, as far as the material treated is concerned, as follows:

TABLE 3

COVARIANCE TESTS FOR THE REGRESSION OF BODY LENGTH ON SCALE LENGTH

TEST	DATA COMPARED	HYPOTHESES	STATISTIC*	F. RATIO* (.05 l. of s.)
1	1938, 1939, 1940, 1941, and 1942 year-classes, ages 1–5, data pooled from all ports.	For all the data, the relationship can be expressed by a straight line.	.19	2.93
		No difference between the regression coefficients of each year-class.	5.76	2.37
2	1938 class in: Pacific Northwest, San Francisco, Monterey, San Pedro	For all the data, the relationship can be expressed by a straight line.	.84	3.67
		No difference between the regression coefficients of the 1938 class in each port.	.30	2.60
3	1939 class in: Pacific Northwest, San Francisco, Monterey, San Pedro	For all the data, the relationship can be expressed by a straight line.	.15	3.67
		No difference between the regression coefficients of the 1939 class in each port.	2.12	2.60
4	1940 class in: Pacific Northwest, San Francisco, Monterey, San Pedro	For all the data, the relationship can be expressed by a straight line.	.09	3.67
		No difference between the regression coefficients of the 1940 class in each port.	2.87	2.60
5	1941 class in: San Francisco, Monterey, San Pedro	For all the data, the relationship can be expressed by a straight line.	.16	5.66
		No difference between the regression coefficients of the 1941 class in each port.	3.47	3.07
6	1942 class in: Pacific Northwest, San Francisco, Monterey, San Pedro	For all the data, the relationship can be expressed by a straight line.	.23	3.67
		No difference between the regression coefficients of the 1942 class in each port.	3.40	2.60

* Statistic to test if the relationship can be expressed by a straight line:

$$\frac{S_1}{N-2p} \cdot \frac{2p-2}{S_2+S_3+S_4} \text{ to be compared with F. ratio: } \frac{N-2p}{2p-2}$$

Statistic to test if there is difference between the regression coefficients of the subclasses involved:

$$\frac{S_2}{p-1} \cdot \frac{N-2p}{S_1} \text{ to be compared with F. ratio: } \frac{p-1}{N-2p}$$

(1) The coefficient of regression of body length on scale length and the body intercept are significantly different for the 1938, 1939, 1940, 1941, and 1942 classes (Test 1).

(2) There is no significant difference among the regression constants for the whole 1938 class regardless of area of capture (Test 2). The same is true for the 1939 class (Test 3). The regression constants for the 1940 class are significantly different for fish sampled in the Pacific Northwest, San Francisco, Monterey, and San Pedro (Test 4). This also is true of the 1941 class (Test 5) and the 1942 class (Test 6).

(3) The relationship between body length and scale length for 1- to 5-ring fish of the year-classes involved does not

TABLE 4
REGRESSION VALUES FOUND FOR HOMOGENEOUS GROUPS

GROUPS	N*	$b_{y.x}$†	s_x‡	s_y§	r‖	$s_{y.x}$¶	$s_{b_{y.x}}$**	Y.††	X.‡‡	a§§	(Y)‖‖	(X)¶¶
1938 year-class......	237	15.90	.58	14.42	.64	11.1	1.24	210	5.20	127	202	124
1939 year-class......	483	21.90	.70	19.89	.77	12.7	.85	211	5.17	98	204	154
1940 year-class in:												
Pacific Northwest.	14	57.00	.17	22.43	.43	20.3	32.54	226	5.76	−102	222	164
San Francisco.....	50	11.66	.54	13.14	.479	11.5	3.01	227	5.65	161	228	170
Monterey........	75	20.45	.73	22.23	.67	16.4	2.60	213	5.30	105	200	152
San Pedro........	70	21.38	.55	14.56	.81	8.5	1.86	201	5.00	94	204	152
1941 year-class in:												
San Francisco.....	5	5.00	.55	10.00	.28	9.6	8.10	220	5.80	191	222	170
Monterey........	26	25.36	.57	20.28	.71	14.3	4.99	210	5.40	73	200	152
San Pedro........	90	16.64	.42	10.94	.64	8.4	2.10	199	5.06	115	196	146
1942 year-class in:												
Pacific Northwest.	7	28.00	.72	24.20	.83	13.5	7.63	233	5.63	75	216	164
San Francisco.....	23	10.50	.58	16.62	.37	15.4	5.65	217	5.53	159	216	164
Monterey........	68	25.40	.73	22.89	.81	13.3	2.20	200	5.13	70	186	140
San Pedro........	51	26.20	.60	18.94	.83	10.5	2.35	202	5.10	68	202	152

* N = number of items in the group.

† $b_{y.x}$ = regression coefficients of y on x = $\frac{N\Sigma X'Y' - (\Sigma X')(\Sigma Y')}{N\Sigma X'^2 - (\Sigma X')^2}$

[For actual calculation the uncoded values were used: X′ = (X′) − 30X; Y′ = (Y) − Y]

‡ s_x = standard deviation of mean scale length = $\frac{1}{5}\sqrt{\frac{\Sigma X'^2 \frac{(\Sigma X')^2}{N}}{N-1}}$

§ s_y = standard deviation of mean body length = $\frac{1}{2}\sqrt{\frac{\Sigma Y'^2 \frac{(\Sigma Y')^2}{N}}{N-1}}$

‖ r = correlation coefficient = $\frac{s_x}{s_y} b_{y.x}$

¶ $s_{y.x}$ = standard error of estimate = $s_y(1-r)$

** $s_{b_{y.x}}$ = variance of regression coefficient = $\frac{s_{y.x}}{s_x\sqrt{N-1}}$

†† Y. = mean body length = $\frac{\Sigma Y'}{N}$ + assumed mean of Y

‡‡ X. = mean scale length = $\frac{\frac{\Sigma X'}{N} + \text{assumed mean of X}}{30}$

§§ a = body intercept = Y. − bX.

‖‖ (Y) = assumed mean of Y

¶¶ (X) = assumed mean of X

deviate significantly from a straight line (Tests 1 to 6 and Table 4).

(4) The values found for the y-intercepts are so great that it seems likely that the regressions do not pass through the origin (0.0) although this was not tested.

These answers imply that, in order to back-calculate lengths, the formula 1n = a + bsn (in which 1n = body length at age n and sn = scale length at age n) should be used in which, for each homogeneous group, the values of "a" and "b" are different (see Table 4). These values are used in a separate work for a study of the rate of deceleration of growth in the Pacific sardine. It should be noted that the formula given above, which applies to averages, suits our purpose quite well as long as our sample is reasonably representative of the commercial catch; also, that its use avoids an error likely to be introduced when only the individual relationship, body length–scale length, is used, i.e., the error caused by the use of scales of different parts of the body.

SUMMARY AND CONCLUSIONS

To determine what method will most accurately serve to back-calculate lengths of pilchard at earlier ages, covariance analysis of observed scale and body-length regressions is useful. Emphasis is given to the most appropriate sampling criteria to be used in order to get the regression values, as well as to the actual method of sampling.

REFERENCES

BLACKBURN, M. 1949. Age, rate of growth, and general life-history of the Australian pilchard (*Sardinops neopilchardus*) in New South Wales waters. *Commonwealth Scientific and Industrial Research Organization Bul.* 242, Div. Fisheries Rpt. 12: 1–86, figs. 1–8, pls. 1–8.

BUCHANAN-WOLLASTON, H. J. 1934. The theory of variation, correlation and regression. Its relevance in researches on proportional growth. *Cons. Perm. Internat. L'Explor. Mer, Rapp. Proc.-Verb.* 89 (4): 33–44, figs. 1–3.

DIXON, W. J., and F. G. MASSEY. 1949. *Introduction to statistical analysis.* Statistical Laboratory, Department of Mathematics, University of Oregon, Eugene. viii+217 pp., 14 tables.

FELIN, F. E., and J. B. PHILLIPS. 1948. Age and length composition of the sardine catch off the Pacific Coast of the United States and Canada. *Calif. Div. Fish and Game, Fish Bul.* 69: 1–122, 1 fig.

KENDALL, M. G. 1946. *The advanced theory of statistics. Vol. II.* Chas. Griffin & Co., Ltd. vii+521 pp., figs. 19.1–30.10.

LEA, EINAR. 1910. On the methods used in the herring-investigations. *Cons. Perm. Internat. L'Explor. Mer, Publ. Circonstance* 53: 7–174, 10 figs., 1 pl.

——— 1913. Further studies concerning the methods of calculating the growth of herrings. *Cons. Perm. Internat. L'Explor. Mer, Publ. Circonstance* 66: 3–36, 6 figs.

LEE, R. M. 1912. An investigation into the methods of growth determination in fishes. *Cons. Perm. Internat. L'Explor. Mer, Publ. Circonstance* 63: 1–34, 3 figs.

OTTESTAD, P. 1938. On the relation between the growth of the fish and the growth of the scales. *Cons. Perm. Internat. L'Explor. Mer, Rapp. Proc.-Verb.* 108 (4): 24–31, 3 figs.

SCHUCK, H. A. 1949. Problems in calculating size of fish at various ages from proportional measurements of fish and scale sizes. *Jour. Wildlife Management* 13 (3): 298–303, 1 fig., 11 pls.

WALFORD, L., and K. MOSHER. 1943*a*. Studies on the Pacific pilchard or sardine (*Sardinops caerulea*) 2. Determination of the age of juveniles by scales and otoliths. *U. S. Fish and Wildlife Serv., Spec. Sci. Rpt.* 20: 1–19, figs. 1–32.

——— and——— 1943*b*. Studies on the Pacific pilchard or sardine (*Sardinops caerulea*) 3. Determination of age of adults by scales, and effect of environment on first year's growth as it bears on age determination. *U. S. Fish and Wildlife Serv., Spec. Sci. Rpt.* 21: 1–22, figs. 1–6.